面向 WSI 的乳腺病理亚型分类研究 *

陈金令†，李　洁，赵成明，刘　鑫
(西南石油大学 电气信息学院，成都 610599)

摘　要：为实现乳腺病理 WSI 图像的精准分类，提出了一种基于混合连接的门控卷积神经网络分类方法。搭建了局部残差连接和全局稠密连接的混合模块，将压缩激活门控单元嵌入混合模块，建立了混合模块与过渡层交替连接的骨干网络。结合基于四叉树分割的图像数据增强方法训练模型，基于 BreastSet 临床数据集的实验结果得出，该方法的图像级、患者级和病理级准确率分别达到 92.24%、92.83%和 92.18%，相较其他方法，准确率提高，参数量和计算量降低，更具有临床应用价值。

关键词：全切片图像；乳腺病理亚型分类；计算机辅助诊断；门控卷积网络；混合连接
中图分类号：TP391.5　　doi: 10.19734/j.issn.1001-3695.2022.03.0087

Research of breast pathological subtype classification on WSI

Chen Jinling†, Li Jie, Zhao Chengming, Liu Xin
(*School of Electrical Information*, *Southwest Petroleum University*, *Chengdu* 610599, *China*)

Abstract: For precise classification of pathological breast WSI images, the study proposed the gated convolution network based on hybrid connection (HC-GCN) , set up a hybrid block of local residual connection and global dense connection, and through embedding the Squeeze -Excitation-and-Gated (SEG) module into the hybrid block, established a backbone network for alternate connection between the hybrid block and the transition layer. In combination with a training model for image data enhancement based on quad-tree image segmentation method, the experimental results based on the BreastSet clinical data set showed 92.24%, 92.83% and 92.18% of accuracy at the image level, patient level and pathology level respectively. Therefore, compared with other methods, this method is of great clinical application value for improved accuracy as well as reduced number of parameters and amount of computation.

Key words: WSI; breast pathological subtype classification; computer auxiliary diagnosis; gated convolutional network; hybrid connection

0　引言

乳腺癌是中国女性第一高发恶性肿瘤，发病率逐年升高，严重威胁中国女性的生命健康[1]。早期乳腺癌是可治愈性疾病，精准地诊断乳腺癌可以最大程度地提高患者的生存概率和生活质量。随着智能算法的发展、医学数据的积累和医疗水平的进步，各种智能算法也逐渐被运用到了医疗领域。传统的图像分类算法由特征提取、特征编码、分类器设计三个步骤组成[2]，人工提取病理特征的过程通常需要占用大量的计算资源，难以实用。卷积神经网络(convolutional neural networks, CNN)具有良好的自动特征建模能力[3]。目前，乳腺病理图像分类领域广泛使用 CNN 构建深度学习(deep learning, DL)模型[4]。

乳腺组织病理图像在临床诊断时发挥着重大作用，乳腺癌病理学图像数据集(breast cancer histopathological database，BreakHis)是已做切分处理的大型公共数据集，目前已有许多研究者围绕其进行病理分类研究。例如，Han[5]等人利用深度学习模型基础层构建 CSDCNN 模型，在多分类任务中获得的平均分类准确率达到 93.3%。Bardou[6]等人采用基于袋模型(Bag of words)、局部约束线性编码器(Locality constrained linear coding)和支持向量机(Support Vector Machine, SVM)分类器的手工特征提取方法，在进行二分类和多分类任务时获得的准确率分别达到 96.15%和 93.31%；第二种是基于数据增强和卷积神经网络设计的深度学习分类方法，二分类和多分类任务获得的准确率分别达到 98.33%和 88.23%。程年[7]等人运用 VGG16 和 InceptionV3 两个深度学习模型，结合迁移学习和数据增强等方法，进行良恶性二分类获得的准确率分别为 94.43%和 82.64%，进行病理亚型八分类获得的准确率分别为 82.97%和 58.78%。Nguyen[8]等人结合调整大小的原始图像等方法建立 CNN 模型，对多种乳腺癌类型进行自动分类，获得的准确率可达 73.68%。Mt[9]等人利用图像增强技术处理每个图像数据，通过注意模块对图像的重要关键区域进行选择和处理，建立 BreastNet 了模型，实验显示良恶性二分类准确率最高达到 98.80%。李雨倩[10]使用深度学习方法作为特征提取器，根据乳腺图像特征构建随机森林模型，该模型进行多分类的准确率达到 92.5%。Jiang[11]等人建立了基于小型 SE-ResNet 模块的卷积神经网络模型，实验显示良恶性二分类和病理亚型八分类的准确率分别为 98.87%~99.34%和 90.66%~93.81%。刘巧利[12]等人提出了基于 DenseNet 网络的乳腺癌分类模型，在二分类任务中得出的识别准确率平均值达到了 99.2%。刘靖雯[13]等人提出了基于 Inception-ResNet-V2 的乳腺癌病理图像识别方法，获得的八分类准确率达到 79.7%。明涛[14]等人结合注意力机制算法，提出基于深度学习的多尺度通道重校准模型，该模型在良恶二分类实验中展示的准确率达到 88.87%。于凌涛[15]等人提出了一种基于卷积神经网络的多任务模型，

收稿日期：2022-03-04；**修回日期：**2022-04-26　**基金项目：**成都市科技厅创新创业资助项目(2018YF0500893GX)
作者简介：陈金令(1975-)，男(通信作者)，河南平顶山人，教授级高级工程师，硕导，博士，主要研究方向为图像处理和光学检测(201899010004@swu.edu.cn)；李洁(1997-)，女，四川德阳人，硕士研究生，主要研究方向为深度学习和医学图像处理；赵成明(1996-)，男，四川泸州人，硕士研究生，主要研究方向为机器学习和图像分割；刘鑫(1996-)，男，四川宜宾人，硕士研究生，主要研究方向为人工智能和红外检测.

模型在二分类任务上达到了 98.55%～99.52%的分类准确率，在多分类上达到了 92.26%～94.85%的分类准确率。赵晓平[16]等人基于稠密卷积神经网络、注意力机制和焦点损失函数提出了乳腺癌组织病理图像的多分类模型，在 BreakHis 数据集上进行验证性实验，得出良恶性二分类识别准确率达到 99.1%，乳腺病理亚型八分类的识别准确率达到了 95.5%。

在对乳腺病理类型进行临床诊断过程中，大多是直接对乳腺病理 WSI(whole slide image, WSI)图像进行直接观察分析，因此，对乳腺病理 WSI 图像数据进行研究更有实用价值。Cruz-Roa[17]等人提出了对于全切片乳腺病理图像中浸润性导管癌组织区的自动检测和视觉分析的深度学习方法，在 49 张 WSI 图像上获得了 84.23%的均衡准确度。Wang[18]等人设计了转移乳腺癌检测器对哨落淋巴结的 WSI 图像进行自动检测，获得的图像分类 AUC 达到 0.995，肿瘤定位得分达到 0.733。Gecer[19]等人建立四个全卷积网络组成的深度学习模型对乳腺 WSI 图像进行五分类，得出的预测效果与 45 位病理学家诊断结果有 55%的相似度。

为实现对乳腺全切片图像病理亚型的精准高效分类，本文利用来自绵阳市中心医院的乳腺病理 WSI 图像，建立 BreastSet 数据集，并提出了基于混合连接的门控卷积网络模型(gated convolution network based on hybrid connection，HC-GCN)对其进行分类研究。模型采用混合连接的组织架构，通过跨层的“短路连接”(shortcuts，skip connection)加速向前传播信息，降低稠密连接导致的庞大冗余性；利用结合了注意力机制的门控结构对特征信息进行有效融合，使得病理类别判别的准确率提高，分类模型的计算复杂度减小。此外，本文针对混合模块特性使用 Tanh′ 函数作为激活函数，避免了 ReLU 函数引起的神经元“坏死”后无法训练问题，从网络结构和激活函数的角度提升了模型性能。

1　基于混合连接的门控卷积网络

1.1　混合连接

He[20]等人提出的残差连接为深度学习突破了发展瓶颈，通过跨层的数据通道结构训练出有效的深层神经网络，建立 ResNet 模型。但实际操作中，ResNet 的网络层数通常较多，所以尽管其模型内部具有参数共享机制，但参数量和计算量仍然较大。另外，残差模块堆叠过程中，早期的特征信息被自然地丢失，太多的残差聚合会导致特征信息损失过多，最终造成无法弥补的误差。DenseNet[21]是由 Huang G 等人提出的以稠密连接为基础结构的网络模型，与 ResNet 同样采用跨层连接的模式，不同的是残差块的输出是输出和输入相加，稠密块的输出是输出和输入在通道维上相联结。DenseNet 减轻了梯度消失问题，促进了特征信息的有效传递和利用，保留了更多的特征信息，对网络深度的要求较低，且每层网络的通道数明显较少，减小了模型的参数量，但稠密连接网络内每一层的特征信息都是前面所有层的聚合，这种做法虽然保护了特征信息的利用，但也导致了模型重复提取冗余信息，使得模型的冗余性极高，计算量居高不下。

为了提高乳腺癌的检测准确率，同时减小算法复杂度和运算时间，本文提出局部的残差连接与全局的稠密连接相聚合的混合连接。混合连接模块的基础结构如图 1 所示。

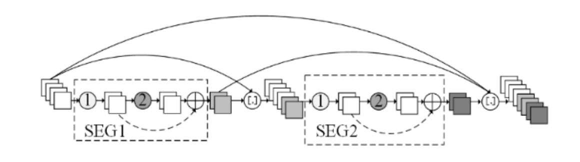


图 1　混合模块

Fig. 1　Hybrid block

图 1 表示由两个 SEG 块进行混合连接组成的混合模块，每一个混合模块都由 2 个或以上的 SEG 块组成。其中，实弧线表示稠密连接，虚弧线表示残差连接，“+”表示两个输入的相加，“[,]”表示通道维上的联结，①和②分别表示 SEG 块中的压缩单元与激活单元。

混合模块的最终目的是代替原本 DenseNet 网络中的稠密块，实现网络模型的混合连接，提高乳腺病理图像病理分类模型的特征提取能力。

1.2　压缩激活门控单元

图 1 中的 SEG1 和 SEG2 表示专为混合模块搭建的特殊结构，功能和原理与瓶颈层(bottleneck layer)类似，称为 SEG 块。SEG 块的具体结构如图 2 所示。

图 2 中，H, W, C 分别表示特征图的高、宽和通道数。

如图 2 所示，在 SEG 模块中，首先，利用卷积和组卷积结构对输入特征图像进行压缩，控制特征维度、减小模型的通道数，进而减小模型所需的参数量和计算量。

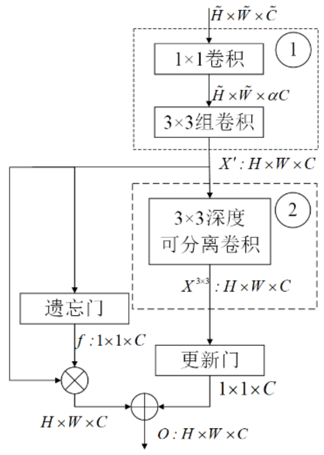


图 2　压缩激活门控单元

Fig. 2　Squeeze -Excitation-and-Gated module (SEG)

其次，运用引入了注意力机制的遗忘门和更新门结构对特征信息进行有效处理，遗忘门使用一个全局上下文模型(global context model)用于计算每个位置的空间注意力特征权重，又加入了 Sigmoid 函数对每个通道进行加权衰减，有效地减小模型所需的计算量；更新门使用 Softmax 函数起到相互比较的作用，选出了最应被注意的特征信息与原信息叠加，促进模型对新特征的挖掘和应用，能够在一定程度上提高模型的分类准确率。同时，更新门和遗忘门的结构既能够像自注意力机制模型一样有效地对全局上下文进行建模，又能够有效地节省计算资源。下面将针对 SEG 块中的压缩与激活、遗忘门和更新门三种结构逐一进行介绍。

1.2.1 压缩与激活

压缩单元的操作步骤主要分为两步：首先，利用 1×1 的标准卷积对输入特征图进行压缩，目的是减小通道数，处理后的通道数将由 $\tilde{C}$ 减小为 αC，其中，α 表示通道的宽度乘数，且 $0<\alpha<1$；然后使用组数为 g，卷积内核为 3×3 的组卷积对特征图进一步压缩，通过设置步长(stride)来实现模型的降采样操作，以减小特征图的高、宽和通道数。两个压缩操作的根本目的都是降低模型的参数量和计算量。

特征信息经过压缩后进入激活单元，运用 3×3 的深度可分离卷积(dethwise convolution)进行激活和数据填充。乳腺病理图像中隐藏的信息相较于常见的分类图像(例如猫狗、鲜花等)隐藏的信息要多，要提取更多的信息，在卷积层需要更多的卷积核，而卷积核的增大会带来计算规模的增大。在输入输出的特征图维度相同情况下，深度可分离卷积所需的输入通道很少、参数和计算的成本极低，参数量和计算量相比与标准卷积显著较少，用深度可分离卷积进行激活和数

据填充可以在不影响效率的情况下提升模型性能。

1.2.2 遗忘门

图 2 中的最左边的连接线和图 1 中的虚线同样表示为 SEG 块中的残差连接，与普通残差连接不同的是，原来的残差连接是通道信息与残差连接信息的简单相加，此处由于同时使用了残差连接和稠密连接，所以为了高效利用在通道中重复传递的特征信息，在残差连接上嵌入了一个引入了注意力机制的遗忘门对重复利用的特征信息进行衰减过滤。

遗忘门为了满足对重用特征的衰减，不仅要保证特征信息的有效流通，而且需将每个通道权重的最终输出映射控制在(0, 1)之间。所以，遗忘门在引用了空间注意力和通道注意力机制的同时，使用了 Sigmoid 函数进行最后的衰减工作。

遗忘门的结构如图 3 所示，操作步骤主要分为以下五步：

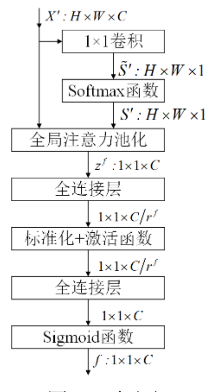


图 3　遗忘门

Fig. 3　Forget gate

Step1：将特征图 $X' \in \mathbb{R}^{H\times W\times C}$ 输入 1×1 的标准卷积，输出通道维数减小的空间注意力特征 $\tilde{S}' \in \mathbb{R}^{H\times W\times 1}$；

Step2：利用 Softmax 函数对空间注意力特征图进行归一化得到新的空间注意力特征 $S' \in \mathbb{R}^{H\times W\times 1}$，其中每个元素可用如下公式表示：

$$S'_{i,j,1} = e \wedge \tilde{S}'_{i,j,1} \Big/ \sum_{x=1}^{H} \sum_{y=1}^{W} e \wedge \tilde{S}'_{i,j,1}$$

其中，$i = 1,\dots,H$，$j = 1,\dots,W$。

Step3：特征图 S' 经过全局注意力池化层(global attention pooling)，生成缩小的全局特征图 $z^f \in \mathbb{R}^{1\times 1\times C}$，第 c 通道全局特征图的表示公式如下：

$$z_c^f = \sum_{x=1}^{H} \sum_{y=1}^{W} X'_{x,y,c} * S'_{x,y,c}$$

其中，$*$ 表示元素乘法。

Step4：全局注意力特征图 z^f 经过两个连续的全连接层。因为网络在训练过程中，上一层训练参数的更新会使后一层输入数据分布也发生变化，两个全连接层之间加入了一个标准化+激活函数层(BN+Tanh)进行缓冲，作用是在两个全连接层中进行一次批量归一化处理操作。

Step5：利用 Sigmoid 函数进行最后的衰减操作。

遗忘门最终的输出特征图表示公式如下：

$$f = \sigma(W_2^f(\mathrm{Tanh}'(BN(W_1^f z^f \mid b_1^f))) \mid b_2^f)$$

其中，f 表示遗忘门的最终输出特征图，W_i^f, b_i^f 分别表示遗忘门中第 i 个全连接层的权重和偏置值；σ 表示 Sigmoid 函数；r^f 表示遗忘门设置的瓶颈比率数；并且 $f \in \mathbb{R}^{1\times 1\times C}$，$W_1^f \in \mathbb{R}^{C\times C/r^f}$，$W_2^f \in \mathbb{R}^{C/r^f\times C}$，$b_1^f \in \mathbb{R}^{C/r^f}$，$b_2^f \in \mathbb{R}^{C}$。

1.2.3 更新门

更新门的设立是为了更有效地处理 3×3 深度可分离卷积得到的新特征，具体结构如图 4 所示。

更新门与遗忘门的差别在于遗忘门的目的是对信息进行衰减，而更新门的目的是促进特征信息的挖掘和使用。二者在结构上也有明显的差异，更新门中没有使用 Sigmoid 函数对特征信息衰减，并且它的最终输出和全局注意力池化层之间有一个简单相加操作。

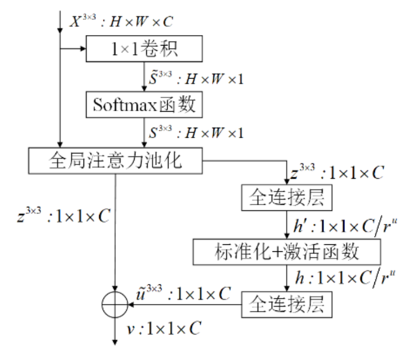


图 4　更新门

Fig. 4　Update gate

令 h 表示标准化+激活函数层后输出的隐藏特征图，其中 $h \in \mathbb{R}^{1\times 1\times C/r^f}$，$r^u$ 表示更新门设置的瓶颈比率数。隐藏特征图的具体公式表示如下：

$$h = \mathrm{Tanh}'(BN(W(z^{3\times 3})) + b)$$

其中，W,b 分别表示第一个全连接层的权重和偏置值，且 $W \in \mathbb{R}^{C\times C/r^u}, b \in \mathbb{R}^{C/r^u}$

隐藏特征图 h 在经过第二个全连接层后生成通道注意力特征图 $\tilde{u}^{3\times 3}$，公式为

$$\tilde{u}^{3\times 3} = W^{3\times 3} h + b^{3\times 3}$$

其中，$W^{3\times 3}, b^{3\times 3}$ 分别表示第二个全连接层的权重和偏置值，且 $W^{3\times 3} \in \mathbb{R}^{C/r^u\times C}, b^{3\times 3} \in \mathbb{R}^{C}$。

第二个全连接层的输出 $\tilde{u}^{3\times 3}$ 与全局注意力池化层的输出 $z^{3\times 3}$ 相加得到更新门的最终输出 v，$v \in \mathbb{R}^{1\times 1\times C}$。

本质上，更新门和遗忘门的作用是将压缩单元形成的紧凑特征图 X' 看做是重用的旧特征，将 3×3 的深度可分离卷积形成的输出 $X^{3\times 3}$ 看做是提取的新特征，最后巧妙地提取和聚合新旧特征构成最终的输出 $O \in \mathbb{R}^{H\times W\times C}$。这种处理方法不仅能够在节省参数量和计算量的同时促进特征的有效重用，而且能够增强对更多特征信息的挖掘能力。

1.2.4 激活函数的选择

激活函数能够为神经网络训练过程加入非线性因素，提高模型的特征表达能力。

图 5 展示了 ReLU、Sigmoid 和 Tanh 三种常用激活函数的函数图像。

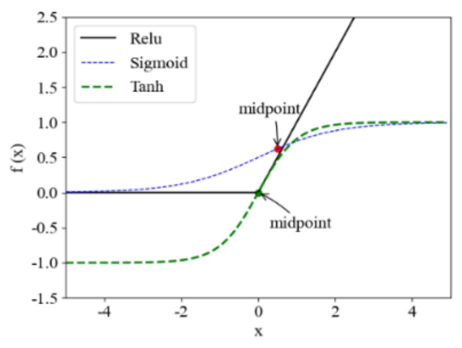


图 5　ReLU、Sigmoid 和 Tanh 三种常用激活函数的函数图像

Fig. 5　The function images of relu, Sigmoid and Tanh

ReLU 函数具有符合神经网络编码特点、克服梯度下降问题、提高计算效率等优点。但从图 5 中的 ReLU 函数曲线可以看出，ReLU 函数有两个致命的缺点。一是当输入神经元 $x<0$ 时，ReLU 的梯度被置零，这很有可能导致 $x<0$ 时神经元“坏死”，输入数据无法激活。二是 ReLU 函数的输出恒为正值，没有负值。因此，ReLU 会在很大程度上破坏模型的特征表示能力，导致模型无法训练出有效的特征信息。

另外，ReLU 函数不会控制数值的幅度，在深层网络中数值幅度变化过大可能导致模型无法训练。Sigmoid 函数和 Tanh 函数能够有效解决神经元“坏死”的问题，保护 $x<0$ 时的特征信息，并且控制数值的幅度。

由图 5 可以看出，采用 Sigmoid 函数作为激活函数时，如果当前输入参数 (ω_0,ω_1) 的最佳优化方向是 $(+d_0,-d_1)$，那么因为 Sigmoid 函数不以 0 为中心，并且输出值恒为正，所以模型将无法实现最快的参数更新，而是走 z 字形逼近最优解。因此，使用以 0 为中心，取值范围为 $[-1,1]$ 的 Tanh 函数可以让模型的收敛速度更快。

另外， Tanh 函数的梯度消失问题比 Sigmoid 函数要轻。但 Tanh 在 $[-1,1]$ 区间上的取值近似线性，这可能会导致提取的图像特征有一定的误差，降低特征学习的准确度。

为了解决 Tanh 函数导致的梯度消失问题和分类误差提高的问题，本文选择 Tanh' 作为 HC-GCN 的激活函数， Tanh' 的具体公式如下：

$$\mathrm{Tanh}' = \mathrm{Tanh}(x)\cdot \mathrm{sigmoid}(x) \tag{1}$$

选用 Tanh' 函数作为激活函数，既可以有效解决“神经元”坏死问题、保护特征信息、控制数值幅度，还能够消除近似线性处理导致的分类误差，获得更好的模型性能和分类结果。另外，混合连接显著加强了梯度的反向传播，可以极大程度地弥补激活函数可能导致的梯度消失的问题。

1.3 模型整体架构

HC-GCN 的整体架构与 DenseNet 相似，不同的是 DenseNet 的主要构建模块是稠密块(dense block)[22]，而 HC-GCN 的主要构建模块是混合模块(hybrid block)。混合模块和稠密块的作用都是定义模型内部瓶颈块的输入输出联结模式。

HC-GCN 模型将需要进行病理类别分类的乳腺病理图像作为网络的输入。首先，输入图像经过一个由 3×3 卷积层、批量标准化(BatchNorm，BN)和 ReLU 函数组成的特征块做标准化处理。然后，传入混合模块与过渡层交互连接的骨干网络提取特征信息，过渡层的主要作用是控制通道数。最后，提取到的特征信息进入全局平均池化(global averagePooling，GAP)层、全连接层(fully connected layer)和 Softmax 分类器，完成最后的分类操作。图 6 中展示的 HC-GCN 的骨干网络是由 3 个混合模块与 2 个过渡层交互连接构成的，其中每个混合模块内部均有多个压缩激活门控单元(Squeeze-Excitation-Gated module，SEG)以稠密连接的模式构成。

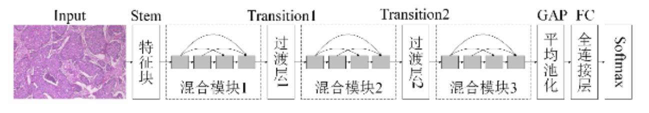


图 6　HC-GCN 整体架构示意图

Fig. 6　Overall architecture schematic of HC-GCN

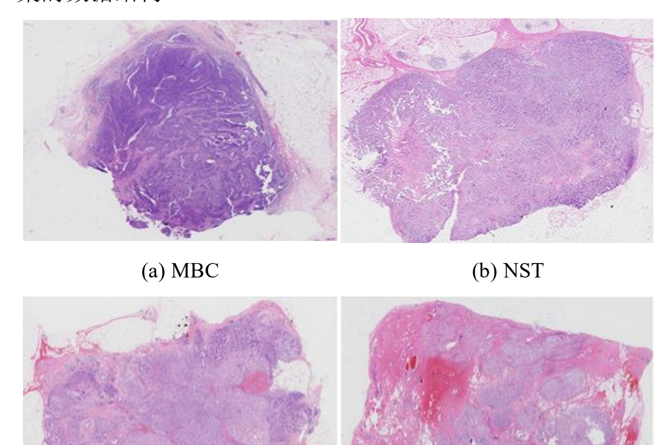
在实际应用中，为了获得更好的乳腺病理图像的病理类别分类效果，可以通过改变网络中的混合模块、SEG 块的数量和调试超参数等方法来构建不同结构的 HC-GCN。

2　数据集来源及处理

2.1　数据集来源

本文采用的图像数据是由绵阳市中心医院提供的乳腺病理 WSI 图像，源文件为 KFB 格式，可利用 K-Viewer 软件对其进行查看和处理。图 7 展示了 BreastSet 数据集中的部分样本。

将原始数据导出格式为 png 的 RGB 三通道图像后可形成分类数据集，数据集包含了来自患不同乳腺癌病症患者的 3498 张的 WSI 病理图像，分类标签共有八种：伴有髓样特征乳腺癌(medullary breast carinoma，MBC)、非特殊类型浸润性癌(non-special type invasive breast carcinoma，NST)、浸润性大汗腺癌(apocrine carinoma，AC)、浸润性微乳头状癌(invasive micropapillary carinoma，IMPC)、浸润性小叶癌(invasive lobular carcinoma，ILC)、黏液癌(mucoid carinoma，MC)、实性乳头状乳腺癌伴浸润(solid papillary caricoma，SPC)、小管癌(tubular carcinoma，TC)。在此将该数据集命名为乳腺病理图像分类数据集(breast pathological image classification data set，BreastSet)。表 1 介绍了 BreastSet 数据集的数据结构。



(a) MBC　(b) NST

(c) IMBC　(d) ILC

图 7　BreastSet 样本数据

Fig. 7　BreastSet sample data

表 1　BreastSet 的数据结构

Tab. 1　Breastset data structure

病理类别	总计	训练集	验证集	测试集
MBC	407	325	41	41
NST	519	415	152	152
AC	133	107	13	13
IMPC	597	477	60	60
ILC	757	607	75	75
MC	464	372	46	46
SPC	496	398	49	49
TC	125	101	12	12
总计	3498	2802	348	348

2.2　数据增强

WSI 图像的尺寸和内存占用较大，因此 BreastSet 数据集中的图像数据难以被计算机直接识别与应用。为了保证图像数据的精准读取，提高模型的特征表达能力，可对数据集中的数据做数据增强处理。除改变尺寸等基础的数据增强方法以外，还可使用如图 8 所示的基于四叉树分割的图像数据增强方法[23]。

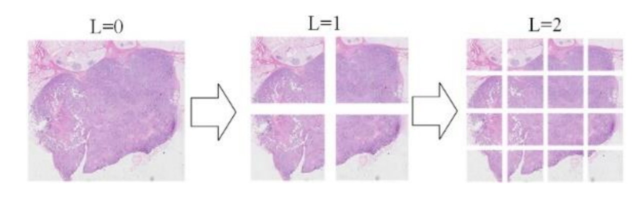


图 8　四叉树图像分割方法

Fig. 8　Quad-tree image segmentation method

四叉树是一种树状数据结构，在每一个节点上会有四个子区块，常应用于二维空间数据的分析与分类。它将数据区分成为四个象限。数据范围可以是方形或矩形或其他任意形状。四叉树有一个连续的结构组成，在每一层，将上一层的输入图像均等分成 4 个部分。在分割之后，每块子图像都被认为与原始图像具有相同的类标签。

2.3　融合算法

由于采用了基于四叉树的图像分割方法，一张图片被分割成多个子图像块，每个图像块在经过模型计算可能产生不同的分类结果，故需要融合算法将每个图像块的分类结果有效地整合起来，常用的融合算法有和规则、乘积规则、最大值规则、多数投票规则等。和规则、乘积规则和最大值规则算法的具体计算公式[24]如下：

和规则：
$$\phi = \arg\max_{k=1}^{K} \sum_{i=1}^{N} p_i(k) \tag{2}$$

乘积规则：
$$\phi = \arg\max_{k=1}^{K} \prod_{i=1}^{N} p_i(k) \tag{3}$$

最大值规则：
$$\phi = \arg\max_{k=1}^{K} \max_{i=1}^{N} p_i(k) \tag{4}$$

其中，$p_i(k)$ 表示为1张图片的第 i 个子图像块被模型分为 类的概率值；K 表示病理类别总数；N 表示 1 张图片的切分块总数。本文采用的是多种融合算法相结合的整合方法。图 9 展示了 1 张病例图片的完整融合分类结果过程，最后输出的分类结果由三种融合规则得出的分类结果投票决定。

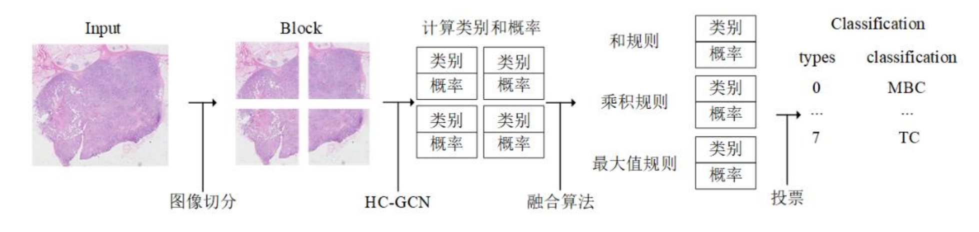


图 9　融合分类结果示意图

Fig. 9　Schematic diagram of fusion classification results

2.4　评价标准

记录模型训练和测试的结果，是进一步完善模型和评价模型质量的依据[25]。为了评价模型在面向 WSI 的乳腺病理亚型分类任务上的适用性，本文计算了分类准确率(accuracy)、模型参数(params)、计算浮点数(FLOPs)、精确率(查准率, precision，P)、召回率(查全率 recall, R)和 F1 评分(F1 scores, F1)等参数进行评估。

在医学领域，计算机辅助诊断模型的分类准确率包含图像级、患者级和病理级三种。图像级的准确率(image-level accuracy，IA)指的是被正确分类的图像数量占样本总数量的比率，患者级的准确率(patient-level accuracy，PA)指的是每位患者对应的乳腺病理图像分类准确率平均值，病理级的准确率(pathological-level accuracy，PLA)指的是多分类情况下的每种病理类型对应的乳腺病理图像分类准确率平均值。

精确率、召回率、F1 评分是衡量模型性能的重要指标。精确率是被诊断为阳性的患者是真阳性的概率，召回率是真阳性患者被诊断为阳性的概率。模型具有较高的精准率可以减少误诊率，防止正常人被错误治疗和医疗资源的浪费；模型具有较高的查全率可以减少漏诊率，防止患者错过最佳治疗时间。

F1 分数(F1 Score)是统计学中用来衡量模型精确度的重要指标[25]，也是模型精准率和召回率的一种调和平均。具体计算公式见式(5)~(7)。

$$\text{precision} = \frac{TP}{TP+FP} \tag{5}$$

$$\text{recall} = \frac{TP}{TP+FN} \tag{6}$$

$$F1 = 2\times\frac{\text{precision}\times\text{recall}}{\text{precision+recall}} \tag{7}$$

其中，TP 表示真阳性数量，FP 表示假阳性数量，FN 表示假阴性数量。

在对病理进行多分类的情况下，计算模型的精确率、召回率和 F1 评分时需要计算相应的宏平均值(Macro)和微平均值(Micro)。宏平均值是指分别计算每个病理类别的精确率、召回率和 F1 评分的平均值，平等地对待每个类别，因此容易受到样本数量较少的类别的影响。微平均值是指不区分类别，整体计算精确率、召回率和 F1 评分，平等的对待每个样本，因此容易受到样本数量较多的类别的影响。根据计算原理，Micro-P，Micro-R 和 Micro-F1 在数值上与 IA 相等。

3　实验

3.1　实验环境

实验用的硬件配置和运行环境如表 2 所示。

表 2　实验环境

Tab. 2　Lab environment

环境	名称及对应版本
操作系统	Ubuntu 18.04.2
编程语言	Python 3.6.13
CPU	Intel Core i7-10870H(2.20GHz)
GPU	NVDIA GeForce GTX 1660Ti
框架	PyTorch1.6, Torchvision0.7.0

3.2　模型架构与超参数设置

为了获得更好的模型分类效果，经过多轮的测试和调整混合模块中的 SEG 块的数量和增长率(growth rate)等结构和参数设置，构建了一个由 3 个混合模块，每个混合模块包含 8 个 SEG 块的 HC-GCN 模型。表 3 展示了模型的具体结构设置。

表 3　HC-GCN 架构配置

Tab. 3　HC-GCN Architecture for breasthis Classification

Stage	HC-GCN2	IR
Stem	[3×3 Conv-BN-ReLU]×3	224×224
	3×3 max pool	112×112
Hybrid Block1	SEG×8，k=24	56×56
transition1	SEG×1	56×56
Hybrid Block2	SEG×8，k=36	28×28
transition2	SEG×1	28×28
Hybrid Block3	SEG×8，k=64	14×14
Classification	global average pool	1×1
	FC，Softmax	-

另外，模型中几个比较重要的超参数设定如下：批样本大小(batch size)为 16, epoch 为 100，学习率(learning rate)为 0.1，过渡层和混合模块之间的压缩因子 θ 为 0.5；SEG 块中 3×3 组卷积的组数 g 为，宽度乘数 α 为 4，更新门和遗忘门的缩减比($r^u = r^f$)为 2，步长 S 为 1；过渡层中设置 g 为 1，$r^u = r^f$ 为 2，α 为 1.5，步长 S 为 2；深度可分离卷积核大小(kernel size)为 2。

3.3　训练与测试

如表 1 所示，实验前按照 8：1：1 的比例分配训练集、验证集和测试集。本文对乳腺病理图像数据增强的第 0 层和第 1 层进行了训练，即训练时分为直接使用原始数据进行训

练和对图像进行四切分后训练两种模式，四切分时每张原始图像被切割成 4 块进行训练和投票融合。

图 10 是 HC-GCN 基于 BreastSet 数据集得出的训练集损失曲线和验证集图像级准确率曲线，图中向上弯曲的实线曲线表示对原数据分类的验证集图像级准确率曲线(Acc0)，向下弯曲的实线曲线表示对原数据分类的训练集损失曲线(Loss0)；向上弯曲的点状曲线表示四切分训练模式下得出的验证集图像级准确率曲线(Acc1)，向下弯曲的点状曲线表示四切分训练模式下得出的训练集损失曲线(Loss1)。

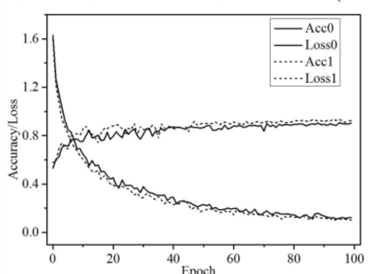


图 10　训练集损失曲线和验证集图像级准确率曲线

Fig. 10　Image-level Accuracy of Verification and Loss of Training

由图 10 可知，HC-GCN 在原数据集和四切分模式上进行训练时都能获得较好的准确率结果和较少的损失，说明本文建立的模型具有一定的准确性和科学性；模型的训练集和验证集损失曲线显示在 BreastSet 数据集上的训练过程较平稳，说明模型具有稳定性；当 epoch 大于等于 40 时，准确率和损失的数值就已经较高并逐渐收敛，说明模型具有收敛性。

对比图 10 中两种训练模式的验证图像级准确率曲线，可以看出两种训练模式下 HC-GCN 的训练准确率曲线平滑，获得的准确率数值较高，说明 HC-GCN 模型适用于对当前乳腺病理数据集进行亚型分类研究。

3.4　实验结果分析

表 4 展示了本文所提方法和其他较经典的深度学习方法基于 BreastSet 数据集的分类实验结果。

从表 4 中各种模型在原始数据训练模式和四切分训练模式下对应得出的数据可以看出，四切分训练模式得出的准确率、F1 评分等准确率性能数据均优于原始数据训练模式。同时，从图 10 可以看出，四切分训练模式相比原始数据训练模式，验证集图像准确率曲线数值普遍更高，训练集损失曲线数值普遍更低，证明使用基于四叉树分割的图像数据增强方法有助于提高模型对于乳腺病理 WSI 图像的分类性能。

HC-GCN 只需要用到很少的参数量和计算量就可以达到较好的病理类型识别精度，有助于降低模型对计算资源的依赖、增加模型的扩展性和可移植性。

3.5　与其他方法比较

本文还选用了多种经典热门的图像分类模型与 HC-GCN 进行对比，每个模型均基于 BreastSet 数据集进行了病理亚型分类，实验结果如表 4 所列。

从表 4 可以看出，许多优秀模型均能够在乳腺病理图像分类任务上获得较高的分类准确率。与各级别准确率最高的网络模型 DenseNet121 相比，HC-GCN 在原始数据训练模式下获得的图像级、患者级和病理级准确率分别提高了 1.14%、0.23%和 0.65%；在四切分训练模式下获得的图像级、患者级、病理级准确率分别提高了 0.37%，0.40%和 0.65%，说明 HC-GCN 相比 DenseNet 模型能够获得更高的分类准确率；并且 HC-GCN 所需的参数量和计算浮点数大约分别只占 DenseNet121 网络模型的 3/8 和 1/6，有效地减轻了特征重用机制导致的冗余度过大问题，节省了计算资源。与具有较小参数量的模型 ShuffleNetv2 相比，HC-GCN 在使用更小的参数量和计算量的情况下，原始数据训练模式下的图像级、患者级和病理级准确率分别提高了 9.77%、9.53%和 9.45%，四切分训练模式下的图像级、患者级和病理级准确率分别提高了 5.57%，3.53%和 17.61%。与其他具有更小的计算量的模型 AlexNet、MobileNetv2 和 MobileNetv3 相比，尽管 HC-GCN 多使用了 0.19GB~0.28GB 的计算量，但原始数据训练模式下的图像级、患者级和病理级别准确率分别提高了 5.46%~19.54%、5.84%~16.89%和 6.66%~29.32%，四切分训练模式下的图像级、患者级和病理级准确率分别提高了 6.24%~12.33%、5.93%~12.37 和 5.24%~11.75%。

表 4　实验结果

Tab. 4　Accuracy, Params, Flops, Precision, Recall and F1 Scores of Classification

模型	L=0						L=1						Params	Flops
	IA	PA	PLA	Macro-P	Macro-R	Macro-F1	IA	PA	PLA	Macro-P	Macro-R	Macro-F1	/MB	/GB
AlexNet	78.74	77.64	77.68	85.05	77.68	81.53	85.14	84.25	83.44	83.55	83.44	83.49	16.63	0.30
VGG11	84.48	86.19	85.75	87.52	85.75	87.03	**86.73**	**87.21**	**85.00**	**85.20**	**85.00**	**85.01**	132.86	7.62
VGG13	**85.06**	**87.10**	**86.33**	**88.20**	86.33	**87.71**	85.26	85.59	81.53	84.07	81.53	82.49	133.05	11.32
VGG16	81.61	82.72	82.51	82.84	**82.51**	82.65	84.28	84.10	81.30	82.71	81.41	81.52	138.36	15.48
VGG19	81.90	83.91	83.65	87.38	83.65	85.25	82.94	83.49	79.47	81.82	79.47	80.18	133.05	19.65
GoogLeNet	84.77	86.78	85.90	88.68	85.90	87.86	88.93	86.80	87.18	88.20	87.17	87.58	7.00	1.59
ResNet34	62.93	65.51	48.45	59.2	48.45	51.96	68.49	69.37	69.74	69.58	70.07	69.82	21.99	6.67
ResNet50	69.83	71.75	59.07	65.13	59.07	61.71	71.78	72.83	71.95	73.12	78.90	75.91	25.56	4.11
ResNet101	72.13	**74.38**	61.61	**72.71**	61.61	65.08	78.41	79.46	**79.18**	80.29	79.31	79.80	44.55	7.83
ResNet152	67.53	68.36	59.04	67.18	59.04	62.24	78.30	78.55	78.35	77.72	78.56	78.34	57.67	9.18
ResNext50	68.39	69.59	54.84	65.87	54.84	58.29	75.76	76.55	75.52	78.04	77.39	77.72	25.03	4.26
ResNext101	**71.55**	71.83	**62.42**	70.18	**62.42**	**65.66**	**78.67**	**79.71**	79.03	**80.56**	**79.57**	**80.06**	88.89	16.48
MobileNetv2	71.26	73.95	61.51	70.99	61.51	64.90	79.91	80.46	80.43	80.40	86.89	83.52	3.51	0.31
MobileNetv3	**85.34**	**85**	**84.17**	**83.96**	**84.17**	**83.82**	**86.00**	**86.90**	**86.94**	**87.29**	**88.73**	**88.00**	5.48	0.22
DenseNet121	**89.66**	**90.61**	**90.18**	**90.73**	**90.18**	**90.53**	**91.87**	**92.43**	**91.53**	**91.46**	**91.53**	**91.47**	7.98	2.87
DenseNet169	85.87	86.09	84.50	83.86	84.50	84.11	88.79	90.31	89.45	90.23	89.45	90.15	14.15	3.40
DenseNet201	89.37	90.16	89.36	90.57	89.36	90.39	91.26	91.79	91.01	90.93	91.01	90.98	20.01	4.34
DenseNet264	87.08	87.42	87.74	84.88	86.53	85.66	89.08	90.16	89.76	90.57	89.76	90.40	28.68	7.79
EfficientNetb0	87.93	89.04	88.17	89.40	88.17	88.57	88.20	88.78	86.44	87.79	86.44	86.93	5.29	0.65
ShuffleNetv2	81.03	81.31	81.38	79.54	81.38	80.88	86.67	89.30	74.57	87.09	74.57	78.69	2.28	1.08
RegNetx1.2gf	87.36	89.47	89.28	89.08	89.28	89.18	90.53	91.61	89.94	89.88	89.94	89.89	46.11	13.61
HC-GCN	90.80	90.84	90.83	90.90	90.83	90.89	92.24	92.83	92.18	91.97	92.18	92.07	3.06	0.50

注：其中，加粗的数值表示在同家族模型中的同指标最优值，灰色底纹的数值表示其他模型的同指标最优值。

与表 4 列出的其他模型相比，HC-GCN 在使用更小的参数量和计算量的同时获得了更优的准确率、精确率、召回率和 F1 评分数值。实验结果说明 HC-GCN 显著提升了对 WSI 乳腺病理图像的亚型分类性能。

综上所述，与其他方法相比，HC-GCN 不仅能够提高乳腺病理图像的病理类别分类准确率，而且可以减少模型所需的参数量和计算量，节省计算资源。

3.6　消融实验

为了测试和验证本文所提方法和模块的有效性，设计了四组互为对照的消融实验。表 5 展示了消融实验中依次添加改进结构模块的情况。

表 5 中，RC 表示残差连接，FG 表示遗忘门，UG 表示更新门。第一组消融实验中每个瓶颈块包含标准正则化层(Batch normalization)、ReLU 激活函数层、1×1 卷积、3×3 组卷积、3×3 深度可分离卷积和 Dropout 层各一个，设计第一组实验的目的是将其与添加各种模块结构后的模型形成对比，验证各结构是否能够有效提升模型的分类性能。第二组至第四组消融实验中，删除了瓶颈模块中的标准正则化层、ReLU 层和 Dropout 层，依次在瓶颈块中添加残差连接、遗忘门和更新门结构，即逐渐形成本文设计的 SEG 块，这样设计的目的是观察各结构对模型性能的影响。

观察表 5 中对应的准确率数值可知，第二组较第一组、第四组较第三组的准确率数值增幅较大，说明残差连接和更新门结构有利于提高模型的分类准确率。观察第二组和第三组实验数据可知，添加遗忘门的操作不会对模型的分类准确率和参数量产生较大影响，同时遗忘门的衰减操作有效地降低了模型的计算量。观察第三组和第四组实验数据可知，更新门只需要占用较少的参数量和计算量就能够使模型的分类准确率获得较大的提升。综上所述，本文所提的混合连接模式和 SEG 块结构均能有效地提升模型性能。

表 5　消融实验

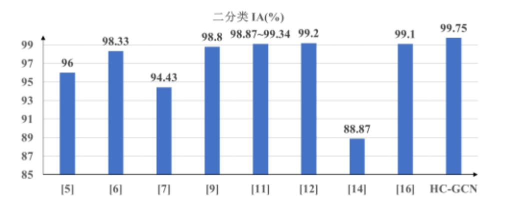
Tab. 5　Ablation experiment

模型	RC	FG	UG	IA	PA	PLA	Params/MB	Flops/GB
1				84.77	84.96	84.22	2.38	2.07
2	√			88.51	89.47	89.29	2.38	0.89
3	√	√		88.79	88.98	88.62	2.72	0.43
4	√	√	√	92.24	92.83	92.18	3.06	0.50

3.7　与其他文献中方法比较

为了证明 HC-GCN 模型的优越性和通用性，本文还将其得出的实验结果与其他文献中使用的乳腺病理图像分类方法进行了对比。

图 11 展示了 HC-GCN 与其他乳腺病理图像分类方法的图像级准确率对比情况，其中的实验数据结果均是基于大型公开的 BreakHis 数据集得出的，[5]~[16]分别表示文献[5]~[16]中所提方法，当文献中所提方法或训练方法不唯一时，图 11 中取文献中展示的图像级准确率最高数值。



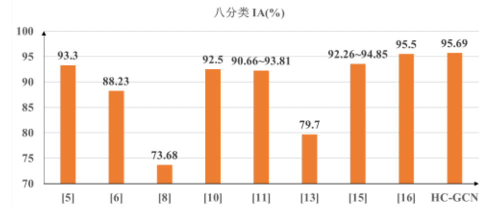


图 11　对比实验结果展示

Fig.11　Compare experimental results

由图 11 可知，在基于 BreakHis 数据集的乳腺病理图像分类实验中，本文提出的 HC-GCN 模型进行良恶性二分类的图像级准确率达到了 99.75%，进行病理亚型八分类的图像级准确率达到了 95.69%，与对比方法相比准确率有所提高。实验结果表明 HC-GCN 模型的分类准确率高于其他文献中的乳腺病理图像分类方法。

另外，图 11 显示本文所提的 HC-GCN 模型在 BreakHis 数据集上也能获得较好的分类效果，说明模型具有良好的通用性和泛化能力。

4　结束语

本文提出基于混合连接的门控卷积网络模型对乳腺病理图像进行亚型分类，运用来自绵阳市中心医院的乳腺病理 WSI 图像建立 BreastSet 数据集，使用基于四叉树的图像分割增强方法进行数据增强，采用融合算法和投票制规则决定每张 WSI 的最终分类。实验结果表明，本文所提的 HC-GCN 模型在乳腺病理 WSI 图像分类问题上能够获得较好的结果，所提的混合连接、更新门和遗忘门结构均能够有效提升模型性能。同时，HC-GCN 模型还具有良好的通用性和泛化能力。

本文方法也适用于其他 WSI 病理图像(例如宫颈癌组织切片图像等)的视觉分类和辅助诊断，能够在一定程度上满足临床需求。在后续的实验中，可考虑对原始 WSI 乳腺病理图像数据进行定位标注，对 BreastSet 数据集进行图像分割和病灶定位，对多种病理类别中难以分类的特征加入自适应调整模块等方面进一步展开研究。

参考文献：

[1] 黄育北，佟仲生，陈可欣，等. 《中国女性乳腺癌筛查指南》解读(精简版) [J]. 中国肿瘤临床，2019, 46 (09): 432-440. (Huang Yubei, Tong Zhongsheng, Chen Kexin, *et al.* Interpretation of guideline for breast cancer screening in Chinese women [J]. Chinese Journal of Clinical Oncology, 2019, 46 (09): 432-440.)

[2] 魏溪含，涂铭，张修鹏. 深度学习与图像识别 [M]. 北京：机械工业出版社，2019: 26-28. (Wei Xihan, Tu Ming, Zhang Xiupeng. Deep learning and image recognition [M]. BeiJing: Machinery Industry Press, 2019: 26-28.)

[3] 马技，李晶皎，李珍妮. 基于视觉注意机制深度强化学习的行人检测方法 [J]. 中国科技论文，2017, 12 (14): 1570-1577. (Ma Ji, Li Jingjiao, Li Zhenni. Deep reinforcement learning with visual attention for pedestrian detection [J]. China Sciencepaper, 2017, 12 (14): 1570-1577.)

[4] 赵凯，安卫超，张晓宇，等. 共享浅层参数多任务学习的脑出血图像分割与分类 [J]. 计算机科学: 2022, 49 (04): 203-208. (Zhao Kai, An Weichao, Zhang Xiaoyu, *et al.* Intracerebral Hemorrhage Image Segmentation and Classification Based on Multi-task Learning of Shared Shallow Parameters [J]. Computer Science, 2022, 49 (04): 203-208.)

[5] Han Z, Wei B, Zheng Y, *et al.* Breast Cancer Multi-classification from Histopathological Images with Structured Deep Learning Model [J]. Scientific Reports, 2017, 7 (1): 4172.

[6] Bardou D, Zhang K, Ahmad S M. Classification of Breast Cancer Based on Histology Images Using Convolutional Neural Networks [J]. IEEE Access, 2018: 1-1.

[7] 程年，俞晨，宁静艳. 基于深度学习网络的乳腺癌图片分类研究 [J]. 软件导刊, 2019, 18 (8): 26-28. (Cheng Nian, Yu Cheng, Ning Jinyan. Classification of Breast Couldcer Images Based on Deep Learning Network [J]. Software Guide, 2019, 18 (8): 26-28.)

[8] Nguyen P T, Nguyen T T, Nguyen N C, *et al.* Multiclass Breast Cancer Classification Using Convolutional Neural Network [C]// Proc of 4th International Symposium on Electrical and Electronics Engineering (ISEE) . New York: IEEE, 2019: 130-134.

[9] Mt A, Kbz B, Be C, *et al.* BreastNet: A novel convolutional neural network model through histopathological images for the diagnosis of breast cancer [J]. Physica A: Statistical Mechanics and its Applications. 2020: 123592.

[10] 李雨倩. 基于深度学习的乳腺组织病理类型的多分类方法研究 [D]. 安徽合肥: 中国科学技术大学, 2019. (Li Yuqian, Research on Multi-classification of Breast Histopathological Patterns based on Deep Learning [D]. Hefei, Anhui: University of Science and Technology of China, 2019.)

[11] Jiang Y, Chen L, Zhang H, *et al.* Breast cancer histopathological image classification using convolutional neural networks with small SE-ResNet module [J]. PloS one, 2019, 14 (3): e0214587.

[12] 刘巧利，闫航，贺鹏飞，杨信志，李彦杰. 基于改进 DenseNet 网络的乳腺癌自动诊断 [J]. 计算机与数字工程, 2019, 47 (10): 2496-2502. (Liu Qiaoli, Yan Hang, He Pengfei, Yang Xinzhi, Li Yanjie. Automatic Diagnosis of Breast Cancer Based on Improved DenseNet Network [J]. Computer & Digital Engineering, 2019, 47 (10): 2496-2502.)

[13] 刘靖雯，黄理灿. 基于 Inception-ResNet-V2 的乳腺癌病理图像识别研究 [J]. 软件导刊, 2020, 19 (5): 225-229. (Liu Jingwen, Huang Lican. Pathological Image Recognition of Breast Cancer Based on Inception-ResNet-V2 [J]. Software Guide, 2020, 19 (5): 225-229)

[14] 明涛，王丹，郭继昌，等. 基于多尺度通道重校准的乳腺癌病理图像分类 [J]. 浙江大学学报（工学版）, 2020, 54 (07): 1289-1297. (Ming Tao, Wang Dan, Guo Jichang, *et al.* Breast cancer histopathological image classification using multi-scale channel squeeze-and-excitation model [J]. Journal of Zhejiang University (Engineering Science) , 2020, 54 (07): 1289-1297.)

[15] 于凌涛，夏永强，王鹏程，等. 基于多任务模型的乳腺癌病理图像分类 [J]. 华中科技大学学报（自然科学版）, 2021, 49 (08): 53-57+69. (Yu Lingtao, Xia Yongqiang, Wang Pengcheng, *et al.* Breast cancer pathological image classification based on multi-task model [J]. J. Huazhong Univ. of Sci. & Tech. (Natural Science Edition) . 2021, 49 (08): 53-57+69.)

[16] 赵晓平，王荣发，孙中波，等. 改进 DenseNet 的乳腺癌病理图像八分类研究 [J]. 计算机工程与应用, 2022, 03 (09): 1-10. (Zhao Xiaoping, Wang Rongfa, Sun Zhongbo, *et al.* Research on Eight Classifications of Breast Cancer Pathological Images Based on Improved DenseNet [J]. Computer Engineering and Applications, 2022, 03 (09): 1-10.)

[17] Cruz-Roa A, Basavanhally A, González F, *et al.* Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks [C]// Medical Imaging 2014: Digital Pathology. San Diego: SPIE, 2014: 904103.

[18] Wang D, Khosla A, Gargeya R, *et al.* Deep learning for identifying metastatic breast cancer [J]. arXiv preprint arXiv: 1606. 05718, 2016.

[19] Gecer B, Aksoy S, Mercan E, *et al.* Detection and classification of cancer in whole slide breast histopathology images using deep convolutional networks [J]. Pattern recognition, 2018, 84: 345-356.

[20] He K, Zhang X, Ren S, *et al.* Deep residual learning for image recognition [C]// Proceedings of the IEEE conference on computer vision and pattern recognition. New York: IEEE, 2016: 770-778.

[21] Huang G, Liu Z, Weinberger K Q, *et al.* Densely connected convolutional networks [C]// Proceedings of the IEEE conference on computer vision and pattern recognition. New York: IEEE, 2017, 4700-4708.

[22] 阿斯顿·张，李沐，扎卡里·C. 立顿，等. 动手学深度学习 [M]. 北京: 中国工信出版集团，人民邮电出版社, 2019: 147-149. (Aston Zhang, Mu Li, Zachary C. Lipton, *et al.* DIVE INTO DEEP LEARNING [M]. BeiJing: China Industry and Information Technology Publishing Group, Posts & Telecom Press, 2019: 147-149.)

[23] Gu J, Wang Z, Kuen J, *et al.* Recent advances in convolutional neural networks [J]. Pattern Recognition, 2018, 77: 354-377.

[24] Yosinski J, Clune J, Bengio Y, *et al.* How transferable are features in deep neural networks? [C]// Advances in Neural Information Processing Systems. Cambridge, MA: MIT Press, 2014: 1-9.

[25] Sokolova M, Japkowicz N, Szpakowicz S. Beyond accuracy, F-score and ROC: a family of discriminant measures for performance evaluation [C]// Australasian joint conference on artificial intelligence. Springer, Berlin: Heidelberg, 2006: 1015-1021.